



## King's Research Portal

DOI:

[10.1016/S1474-4422\(10\)70197-6](https://doi.org/10.1016/S1474-4422(10)70197-6)

*Document Version*

Publisher's PDF, also known as Version of record

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Shatunov, A., Mok, K., Newhouse, S., Weale, M. E., Smith, B., Vance, C., Johnson, L., Veldink, J. H., van Es, M. A., van den Berg, L. H., Robberecht, W., Van Damme, P., Hardiman, O., Farmer, A. E., Lewis, C. M., Butler, A. W., Abel, O., Andersen, P. M., Fogh, I., ... Al-Chalabi, A. (2010). Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study. *Lancet Neurology*, 9(10), 986-94. [https://doi.org/10.1016/S1474-4422\(10\)70197-6](https://doi.org/10.1016/S1474-4422(10)70197-6)

### Citing this paper

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### General rights

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### Take down policy

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# Chromosome 9p21 in sporadic amyotrophic lateral sclerosis in the UK and seven other countries: a genome-wide association study

Aleksey Shatunov, Kin Mok, Stephen Newhouse, Michael E Weale, Bradley Smith, Caroline Vance, Lauren Johnson, Jan H Veldink, Michael A van Es, Leonard H van den Berg, Wim Robberecht, Philip Van Damme, Orla Hardiman, Anne E Farmer, Cathryn M Lewis, Amy W Butler, Olubunmi Abel, Peter M Andersen, Isabella Fogh, Vincenzo Silani, Adriano Chiò, Bryan J Traynor, Judith Melki, Vincent Meininger, John E Landers, Peter McGuffin, Jonathan D Glass, Hardev Pall, P Nigel Leigh, John Hardy, Robert H Brown Jr, John F Powell, Richard W Orrell, Karen E Morrison, Pamela J Shaw, Christopher E Shaw, Ammar Al-Chalabi

## Summary

**Background** Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of motor neurons that results in progressive weakness and death from respiratory failure, commonly within about 3 years. Previous studies have shown association of a locus on chromosome 9p with ALS and linkage with ALS–frontotemporal dementia. We aimed to test whether this genomic region is also associated with ALS in an independent set of UK samples, and to identify risk factors associated with ALS in a further genome-wide association study that combined data from the independent analysis with those from other countries.

**Methods** We collected samples from patients with sporadic ALS from 20 UK hospitals and obtained UK control samples from the control groups of the Depression Case Control study, the Bipolar Affective Case Control Study, and the British 1958 birth cohort DNA collection. Genotyping of DNA in this independent analysis was done with Illumina HumanHap550 BeadChips. We then undertook a joint genome-wide analysis that combined data from the independent set with published data from the UK, USA, Netherlands, Ireland, Italy, France, Sweden, and Belgium. The threshold for significance was  $p=0.05$  in the independent analysis, because we were interested in replicating a small number of previously reported associations, whereas the Bonferroni-corrected threshold for significance in the joint analysis was  $p=2.20 \times 10^{-7}$ .

**Findings** After quality control, samples were available from 599 patients and 4144 control individuals in the independent set. In this analysis, two single nucleotide polymorphisms in a locus on chromosome 9p21.2 were associated with ALS: rs3849942 ( $p=2.22 \times 10^{-6}$ ; odds ratio [OR] 1.39, 95% CI 1.21–1.59) and rs2814707 ( $p=3.32 \times 10^{-6}$ ; 1.38, 1.20–1.58). In the joint analysis, which included samples from 4312 patients with ALS and 8425 control individuals, rs3849942 ( $p=4.64 \times 10^{-10}$ ; OR 1.22, 95% CI 1.15–1.30) and rs2814707 ( $p=4.72 \times 10^{-10}$ ; 1.22, 1.15–1.30) were associated with ALS.

**Interpretation** We have found strong evidence of a genetic association of two single nucleotide polymorphisms on chromosome 9 with sporadic ALS, in line with findings from previous independent GWAS of ALS and linkage studies of ALS–frontotemporal dementia. Our findings together with these earlier findings suggest that genetic variation at this locus on chromosome 9 causes sporadic ALS and familial ALS–frontotemporal dementia. Resequencing studies and then functional analysis should be done to identify the defective gene.

**Funding** ALS Therapy Alliance, the Angel Fund, the Medical Research Council, the Motor Neurone Disease Association of Great Britain and Northern Ireland, the Wellcome Trust, and the National Institute for Health Research Dementias and Neurodegenerative Diseases Research Network (DeNDroN).

## Introduction

Amyotrophic lateral sclerosis (ALS) is a neurodegenerative disease of motor neurons that results in progressive weakness and death from respiratory failure, commonly within about 3 years. About one in 300 people<sup>1</sup> develops ALS but its prevalence is low—five per 100 000 people—because of its poor prognosis. About 5% of patients have a family history of ALS, and occasionally frontotemporal dementia. ALS and frontotemporal dementia can coexist in the same family or individual.

Several studies have reported linkage of ALS–frontotemporal dementia with a locus on chromosome 9p, and specifically with a shared region that spans about 3.6 Mb.<sup>2–7</sup> A genome-wide association study (GWAS) reported that sporadic ALS was associated with single nucleotide polymorphisms (SNPs) in the same region of chromosome 9p.<sup>8</sup> Therefore, this locus is the focus of research into the genetic basis of familial ALS, the relation between ALS and frontotemporal dementia, and the mechanisms underlying sporadic ALS and frontotemporal dementia.

Lancet Neurol 2010; 9: 986–94

This online publication has been corrected. The corrected version first appeared at [thelancet.com/neurology](http://thelancet.com/neurology) on October 13, 2010 with further corrections published on February 21, 2011

Published Online  
August 31, 2010  
DOI:10.1016/S1474-4422(10)70197-6

See [Reflection and Reaction](#) page 945

See [Articles](#) page 978

King's College London, Medical Research Council Centre for Neurodegeneration Research, Department of Clinical Neuroscience (A Shatunov PhD, S Newhouse PhD, B Smith PhD, C Vance PhD, L Johnson PhD, O Abel MSc, I Fogh PhD, Prof P N Leigh FMedSci, Prof C E Shaw MD, Prof A Al-Chalabi PhD), Department of Medical and Molecular Genetics, Guy's Hospital (M E Weale PhD, Prof C M Lewis), Medical Research Council Centre for Social, Genetic and Developmental Psychiatry (Prof A E Farmer FRCPsych, Prof C M Lewis PhD, A W Butler PhD, Prof P McGuffin FRCPsych), and Department of Neuroscience (J F Powell DPhil), Institute of Psychiatry, London, UK; University College London Institute of Neurology, Queen Square, London, UK (K Mok MSc, Prof J Hardy PhD, R W Orrell MD); Department of Neurology, Rudolf Magnus Institute of Neuroscience, University Medical Centre Utrecht, Netherlands (J H Veldink MD, M A van Es MD, Prof L H van den Berg MD); Department of Neurology,

Previous GWAS in ALS have reported association results that have been replicated within the same study but rarely in independent studies.<sup>9</sup> For example, variation in the *ITPR2* gene was associated with ALS in a GWAS and was replicated within the study<sup>10</sup> but has not been found by other investigators.<sup>8</sup> Variation in *DPP6* was reported to be associated with ALS and was replicated within the same study,<sup>11</sup> and a second study found that the same SNPs were the most strongly associated, although they did not achieve genome-wide significance.<sup>12</sup> A subsequent study failed to confirm the findings.<sup>13</sup> Variation in the *UNC13A* gene was strongly associated with ALS and was replicated within the same study,<sup>8</sup> but no successful replication of the association with this locus has been reported.

We therefore aimed to examine previously reported genetic associations for ALS in a GWAS in an independent set of UK samples. We also aimed to combine these data with existing genotypic data from seven other countries to generate a large GWAS in ALS.

## Methods

### Samples

Adults attending one of 20 UK hospitals and who had been diagnosed with ALS by two consultant neurologists were invited to participate in the UK National DNA Bank for Motor Neuron Disease Research (MND DNA bank), a national collection of ALS DNA samples derived from whole blood. Collection began in 2004, and enrolment is ongoing. Patients had no family history of ALS at the time of sampling, were of self-reported white European ancestry, and had onset of weakness on or after January, 2002. Patients provided ethically approved written consent for blood to be drawn. Control samples for the independent analysis were obtained from two sources: the control groups of the Depression Case Control (DeCC) study<sup>14</sup> and the Bipolar Affective Case Control Study (BACCS)<sup>15</sup> and data deposited by Panos Deloukas from the Wellcome Trust Sanger Institute (Cambridge, UK) and published online from the British 1958 birth cohort DNA collection.

Case and control samples for the joint analysis were from the UK, USA, Netherlands, Belgium, France, Italy, Ireland, and Sweden.<sup>8,11,16,17</sup> Demographic, ascertainment, and quality control details of these samples have been published previously.<sup>8,11,16,17</sup> All participating individuals gave written informed consent, and patients were selected for sporadic ALS. Samples were matched for geographical origin and ethnic origin only.

This research was approved by the research ethics committees of each relevant institution for all cohorts.

### Procedures

DNA was extracted by use of standard methods at three centres within 1 week of the blood being drawn (usually on the same day) and was stored centrally at the UK DNA banking network in Manchester. We used a barcode-based sample tracking system to minimise the risk of clerical error.

Genotyping of DNA from samples held with the UK MND DNA bank was done with Illumina HumanHap550 BeadChips (Illumina, CA, USA) by UCL Genomics, London, UK. All DNA samples went through stringent quality control, and processing was done under full laboratory information management system control. The raw data were analysed with GenomeStudio (Illumina) and extracted for statistical analysis.

Quality control procedures were applied to individual and SNP data (webappendix pp 1–3). Individuals were excluded if more than 1% of their genotypic data was missing; if they had abnormal heterozygosity, poor cluster separation, or a sex assignment that conflicted with phenotypic data; if they shared more than 5% of alleles identical by descent with any other study member; or if they were of non-European ancestry. SNPs with minor allele frequency less than 1%, showing non-random missingness between genotypes or cases and controls, or showing departure from Hardy-Weinberg equilibrium ( $p < 1 \times 10^{-6}$ ) were excluded.

For the joint analysis, raw genotypes were obtained from public databases or directly from research groups and merged into a single file for analysis, with a covariate file listing the site of origin. Genotyping and quality control measures for the studies in the joint analysis have been published previously.<sup>8,11,16,17</sup>

All data were deposited at the ALS online genetics database<sup>18</sup> and the European genome-phenome archive.

### Statistical analysis

We did power analysis with the Genetic Power Calculator.<sup>19</sup> For the independent analysis there was 80% power to detect a typical previously reported association with an odds ratio (OR) of 1.2 and minor allele frequency of 0.25 at  $p=0.05$ . We report power at  $p=0.05$  in the independent analysis because, although we have done a GWAS, we are interested in replicating association at five specified loci (*FGGY*, *DPP6*, *ITPR2*, *UNC13A*, and chromosome 9p21.2). For the joint analysis, we investigated the possibility of new associations within the genome and therefore we corrected for multiple testing. We used Bonferroni correction, which assumes the tests of association at each of the 227475 SNPs are independent; the corrected threshold is therefore  $2 \cdot 20 \times 10^{-7}$  ( $0.05/227475$ ), and there was 83.5% power to detect a similar variant at the threshold of  $p=2 \cdot 20 \times 10^{-7}$ . Our aim was to use all of the available samples.

For the joint analysis, population stratification was controlled for by principal component analysis in EIGENSTRAT.<sup>20</sup> We used the Tracy-Widom distribution to calculate the number of significant principal components needed as covariates in subsequent analyses, using the twstats program of the Eigensoft package.<sup>21</sup>

We tested association of genotypes with ALS by logistic regression with 30 principal components as covariates in the program PLINK (version 1.07).<sup>22</sup> We used the expectation-maximisation algorithm—as used in PLINK—for haplotype estimation. Haplotype association with ALS

University Hospital Leuven, and Laboratory for Neurobiology, Vesalius Research Center, VIB, University of Leuven, Leuven, Belgium (Prof W Robberecht MD, P Van Damme MD); Department of Neurology, Beaumont Hospital and Trinity College Dublin, Ireland (Prof O Hardiman MD); Department of Clinical Neuroscience, Umeå University, Sweden (Prof P M Andersen MD); Department of Neurology and Laboratory of Neuroscience, "Dino Ferrari" Centre, Università degli Studi di Milano—IRCCS Istituto Auxologico Italiano, Milano, Italy (Prof V Silani MD); Department of Neuroscience, University of Turin, and Azienda Ospedaliera Universitaria San Giovanni Battista, Turin, Italy (A Chiò MD); Neuromuscular Diseases Research Group, Laboratory of Neurogenetics, National Institute on Aging, National Institutes of Health, Bethesda, USA (B J Traynor MD); INSERM UMR-788 and University of Paris 11, Bicêtre Hospital, Paris, France (Prof J Melki MD); Assistance Publique—Hôpitaux de Paris, Université Pierre et Marie Curie, Hôpital de la Salpêtrière, Paris, France (Prof V Meininger MD); Department of Neurology, University of Massachusetts Medical School, Worcester, MA, USA (J E Landers PhD, Prof R H Brown Jr DPhil); Department of Neurology, Emory University, Atlanta, GA, USA (J D Glass MD); School of Clinical and Experimental Medicine, College of Medicine and Dentistry, University of Birmingham, and Neurosciences Division, University Hospitals Birmingham NHS Foundation Trust, Birmingham, UK (H Pall MD, Prof K E Morrison DPhil); and Academic Neurology Unit, Department of Neuroscience, Faculty of Medicine, Dentistry and Health, University of Sheffield, Sheffield, UK (Prof P J Shaw MD)

Correspondence to: Prof Ammar Al-Chalabi, King's College London, Medical Research Council Centre for Neurodegeneration Research, Institute of Psychiatry, P041, London SE5 8AF, UK. ammar.al-chalabi@kcl.ac.uk

For the British 1958 birth cohort see <http://www.b58cgene.sgu.ac.uk>

For the **MND DNA bank** see  
http://www.mndassociation.  
org/dna\_bank/

For the **amyotrophic lateral  
sclerosis online genetics  
database** see http://alsod.iop.  
kcl.ac.uk

For the **European  
genome-phenome archive** see  
http://www.ebi.ac.uk/ega

For the **1000 Genomes  
Project** see http://  
www.1000genomes.org

See **Online** for webappendix

was tested with a sliding window of varying sizes, including a window sufficient to include all SNPs in the haplotype block that contained the most strongly associated SNPs.

We imputed genotypes by use of Impute2,<sup>23</sup> with data from the 1000 Genomes Project and the International HapMap as templates.<sup>24</sup> Genotype imputation allows the prediction of genotypes at untyped SNPs by the correlation between SNP genotypes in different populations. Association analysis can then be done on the predicted genotypes. Association analysis was done with SNPTEST,<sup>25</sup> which accounts for uncertainty in the imputed genotypes, using the first ten principal components as covariates (the maximum possible). The correction for population stratification in the imputed data will have been weakened by the use of fewer principal components; however, the use of Impute2 increased our ability to deal with genotype uncertainty. We did this to identify variants for follow-up genotyping in a focused region, and thus the trade-off was deemed worthwhile.

Population attributable risk was estimated as:

$$\frac{(p^2 + 2pqg_{AB} + q^2g_{BB}) - 1}{p^2 + 2pqg_{AB} + q^2g_{BB}}$$

where  $q$  is the risk allele frequency,  $p=1-q$ , and genotype relative risk is denoted by  $g$ , with the subscript showing heterozygote or homozygote status. Penetrance for each genotype was estimated as:

$$K \frac{f_{case}}{f_{control}}$$

where  $K$  is the lifetime prevalence and  $f$  is the genotype frequency.

### Role of the funding source

The sponsor of the study had no role in study design, data analysis, data interpretation, writing of the report, or the decision to submit the paper for publication. The Motor Neurone Disease Association is the sponsor of

the UK National MND DNA bank, and therefore had a role in data collection. The corresponding author had full access to all of the data in the study and had final responsibility for the decision to submit the paper for publication.

## Results

For the independent analysis, genotypes were obtained from 663 individuals in the UK MND DNA bank, 1589 control samples from the DeCC study and BACCS, and 2930 control samples from the 1958 birth cohort. After quality control measures, samples were available from 599 people with ALS and 4144 control individuals. Table 1 summarises clinical features of the ALS cohort in the independent analysis.

For the joint analysis, genotypes were obtained from 4312 patients with ALS and 8425 control individuals. After quality control measures, samples were available from 4133 patients with ALS and 8130 control individuals, which were typed for 227475 SNPs to generate 2789513662 genotypes. Table 2 and the webappendix p 4 summarise demographic data.

In the independent analysis, SNPs rs3849942 ( $p=2.22 \times 10^{-6}$ ; OR 1.39, 95% CI 1.21–1.59) and rs2814707 ( $p=3.32 \times 10^{-6}$ ; 1.38, 1.20–1.58) on chromosome 9 showed association with ALS (figure 1, table 3, webappendix pp 5–6). rs903603, which had strong linkage disequilibrium with the associated SNPs ( $D'=0.98$ ), was the SNP most strongly associated with ALS in the independent set ( $p=8.92 \times 10^{-8}$ ; OR 0.71, 95% CI 0.63–0.81). No other loci achieved genome-wide significance.

In the joint analysis, rs3849942 ( $p=4.64 \times 10^{-10}$ ; OR 1.22, 95% CI 1.15–1.30), rs2814707 ( $p=4.72 \times 10^{-10}$ ; 1.22, 1.15–1.30), and rs903603 ( $p=1.37 \times 10^{-7}$ ; 0.86, 0.82–0.91) showed association with ALS (figure 2, table 3, webappendix pp 7–8). SNPs at other loci that were associated with ALS in previous studies did not reach significance in the joint analysis, except the previously associated SNP in *UNC13A* (rs12608932; table 3). The genotype relative risk for rs3849942 was 1.24 (heterozygote) and 1.49 (homozygote), with minor allele frequencies of 0.27 in patients with ALS and 0.24 in controls. The population attributable risk of rs3849942 was 0.10 and the penetrance of the risk allele was 0.004 (heterozygote) and 0.005 (homozygote).

A haplotypic association of rs3849942 and rs10122902 (omnibus  $p=1.51 \times 10^{-11}$ ) was identified in the joint analysis. Of the three haplotypes in the population (GA, AG, and GG), one had equal frequency in people with ALS and controls (GA: frequency cases 0.21, controls 0.20;  $p=0.123$ ), one was associated with increased risk of ALS (AG: frequency cases 0.27, controls 0.23,  $p=7.76 \times 10^{-10}$ ), and was one associated with reduced risk of ALS (GG: frequency cases 0.52, controls 0.57;  $p=4.83 \times 10^{-11}$ ). Similar results were obtained for the independent study: rs3849942 and rs10122902 showed

Case cohort (n=599)

Presentation*	
Bulbar symptoms	165 (28%)
Limb weakness	375 (63%)
Respiratory	11 (2%)
Mixed†	42 (7%)
Age of onset‡ (years)	60.7 (61.7, 22.9–86.5)
Diagnostic delay§ (months)	21.5 (18.3, 2.0–80.1)

Data are number (%) or mean (median, range). \*Data missing for six patients.

†Simultaneously reported weakness of limbs and bulbar musculature. ‡Data missing for three patients. §Time between symptom onset and diagnosis. Data missing for eight patients.

**Table 1: Clinical features of patients with amyotrophic lateral sclerosis in the independent analysis**

	Cases	Controls	Total	Men
Belgium <sup>8</sup>	286	307	593	313
France <sup>16</sup>	230	700	930	805
Netherlands <sup>8,10,12</sup>	977	988	1965	1152
Ireland <sup>11</sup>	210	200	410	219
Italy <sup>17</sup>	259	241	500	271
Sweden <sup>8</sup>	422	441	863	481
USA <sup>16</sup>	923	904	1827	1061
UK <sup>16</sup>	227	205	432	264
UK MND DNA bank	599	0	599	382
DeCC study and BACCS	0	1505	1505	567
1958 birth cohort	0	2639	2639	1357
Total	4133	8130	12 263	6872

The final three rows refer to samples used in the independent study.  
DeCC=Depression Case Control. BACCS=Bipolar Affective Case Control Study.

**Table 2: Characteristics of the component populations for the joint analysis after quality control**

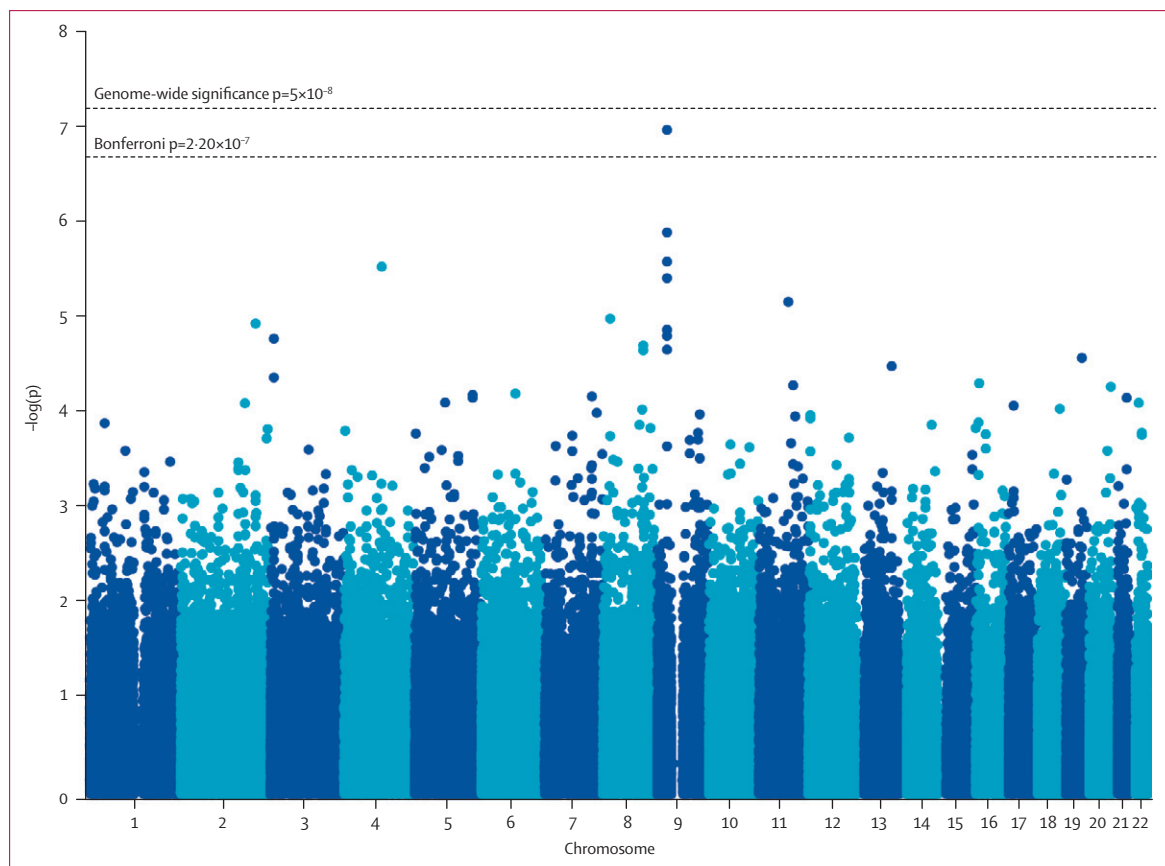
haplotypic association (omnibus  $p=1.01 \times 10^{-7}$ ), and the same three haplotypes had similar control frequencies (GA: frequency cases 0.21, controls 0.19,  $p=0.074$ ; AG: frequency cases 0.30, controls 0.24,  $p=3.18 \times 10^{-6}$ ; GG: frequency cases 0.48, controls 0.57,  $p=4.24 \times 10^{-8}$ ).

Longer SNP haplotypes did not result in stronger association in either part of the study (data not shown). The sliding window showed that pairs of SNPs within a 106.5 kb block of linkage disequilibrium flanked by rs4879515 and rs702231 were nearly all associated with ALS at  $p < 1 \times 10^{-3}$ , whereas outside this region there were no significant associations.

Imputation in the joint analysis in the region of interest (ie, the 3.6 Mb region identified by linkage) gave results for 21899 imputed SNPs that were derived from 519 typed SNPs. Five SNPs were more strongly associated with ALS than was rs3849942; the most strongly associated SNP was a rare variant, chr9:27573575 (minor allele frequency 0.054;  $p=6.40 \times 10^{-15}$ ). The most strongly associated variant with a minor allele frequency greater than 0.1 was rs13295103 ( $p=1.97 \times 10^{-12}$ ; figure 3).

## Discussion

In our independent study we have confirmed association of a chromosome 9 locus with ALS. We did not replicate any of the other previously reported associations, suggesting that our study gave a false-negative result or that they were population specific or false-positive results (panel, table 3).<sup>8-12</sup> In our joint analysis, which is the largest GWAS of ALS to date, chromosome 9p21.2 and



**Figure 1: p values for association in the independent genome-wide association study**



	n for associated SNP in previous studies*	Associated SNP	Controls	p value in previous studies	Independent analysis†			Joint analysis‡		
					p value	Minor allele frequency		p value	Minor allele frequency	
						Cases	Controls		Cases	Controls
FGGY <sup>a</sup>	1287	rs6700125	1567	6.0×10 <sup>-4</sup>	0.71	0.342	0.337	0.08	0.351	0.339
DPP6 <sup>12</sup>	1767	rs10260404	1916	5.4×10 <sup>-8</sup>	0.89	0.397	0.395	0.08	0.391	0.380
ITPR2 <sup>10</sup>	1337	rs2306677	1356	3.28×10 <sup>-6</sup>	0.98	0.100	0.100	0.97	0.100	0.096
UNC13A <sup>8</sup>	4855	rs12608932	14953	2.53×10 <sup>-14</sup>	0.42	0.366	0.354	5.14×10 <sup>-8</sup>	0.382	0.344
9p21.2 <sup>8</sup>	4855	rs2814707	14953	7.45×10 <sup>-9</sup>	3.32×10 <sup>-6</sup>	0.305	0.243	4.72×10 <sup>-10</sup>	0.272	0.236
9p21.2 <sup>8</sup>	4855	rs3849942	14953	1.01×10 <sup>-8</sup>	2.22×10 <sup>-6</sup>	0.304	0.241	4.64×10 <sup>-10</sup>	0.271	0.235
9p21.2 <sup>8</sup>	4855	rs903603	14953	..	8.92×10 <sup>-8</sup>	0.411	0.493	1.37×10 <sup>-7</sup>	0.456	0.492

Previously reported significant associations and the results we obtained for those genes in the two parts of this study are shown.<sup>8-10,12</sup> \*Total combined sample for the reported SNP, which includes a GWAS component and a non-GWAS follow-up study. †599 cases and 4144 controls. ‡4133 cases and 8130 controls.

**Table 3: Associations for genes reported to be associated with amyotrophic lateral sclerosis in previous genome-wide association studies**

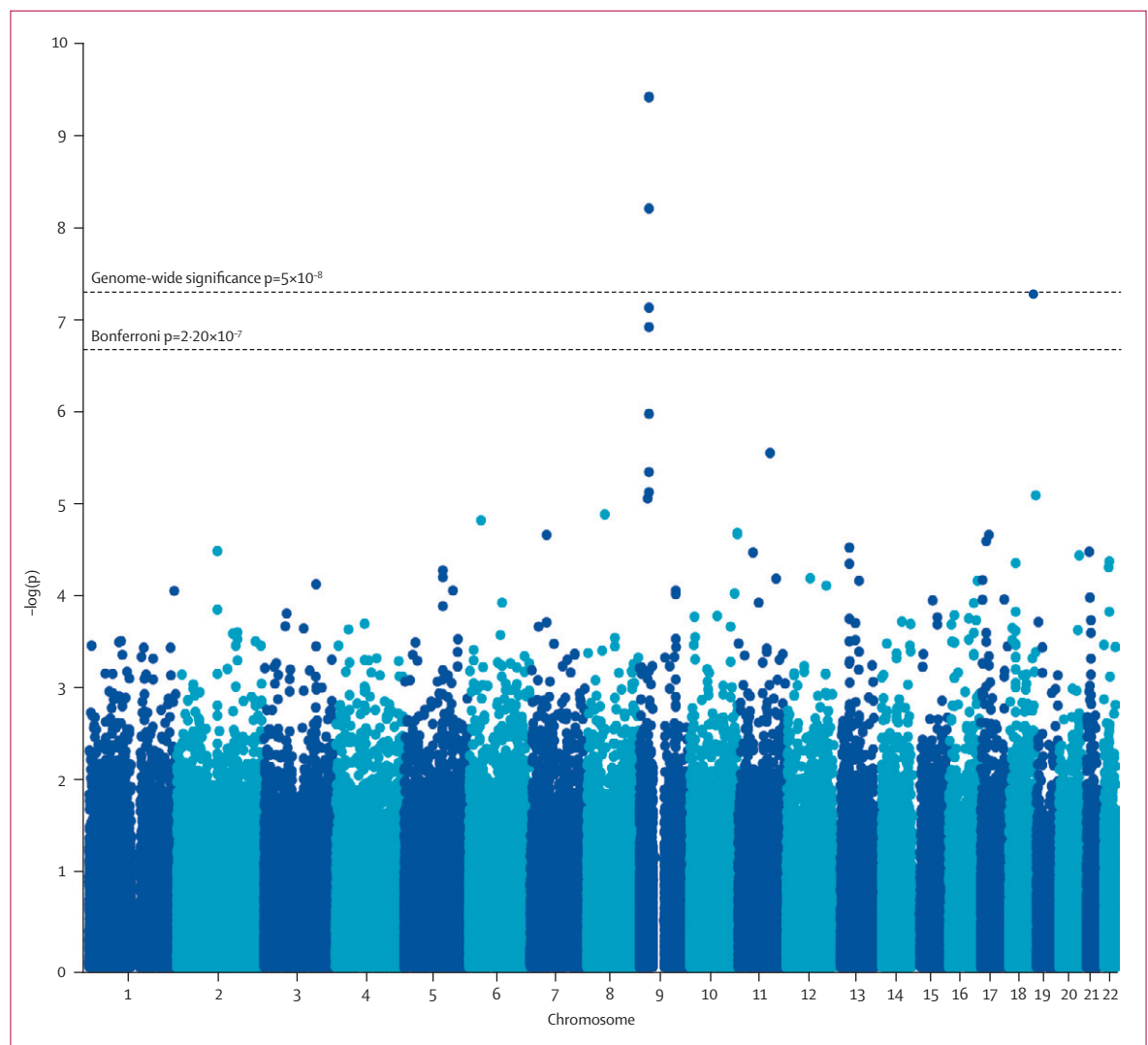
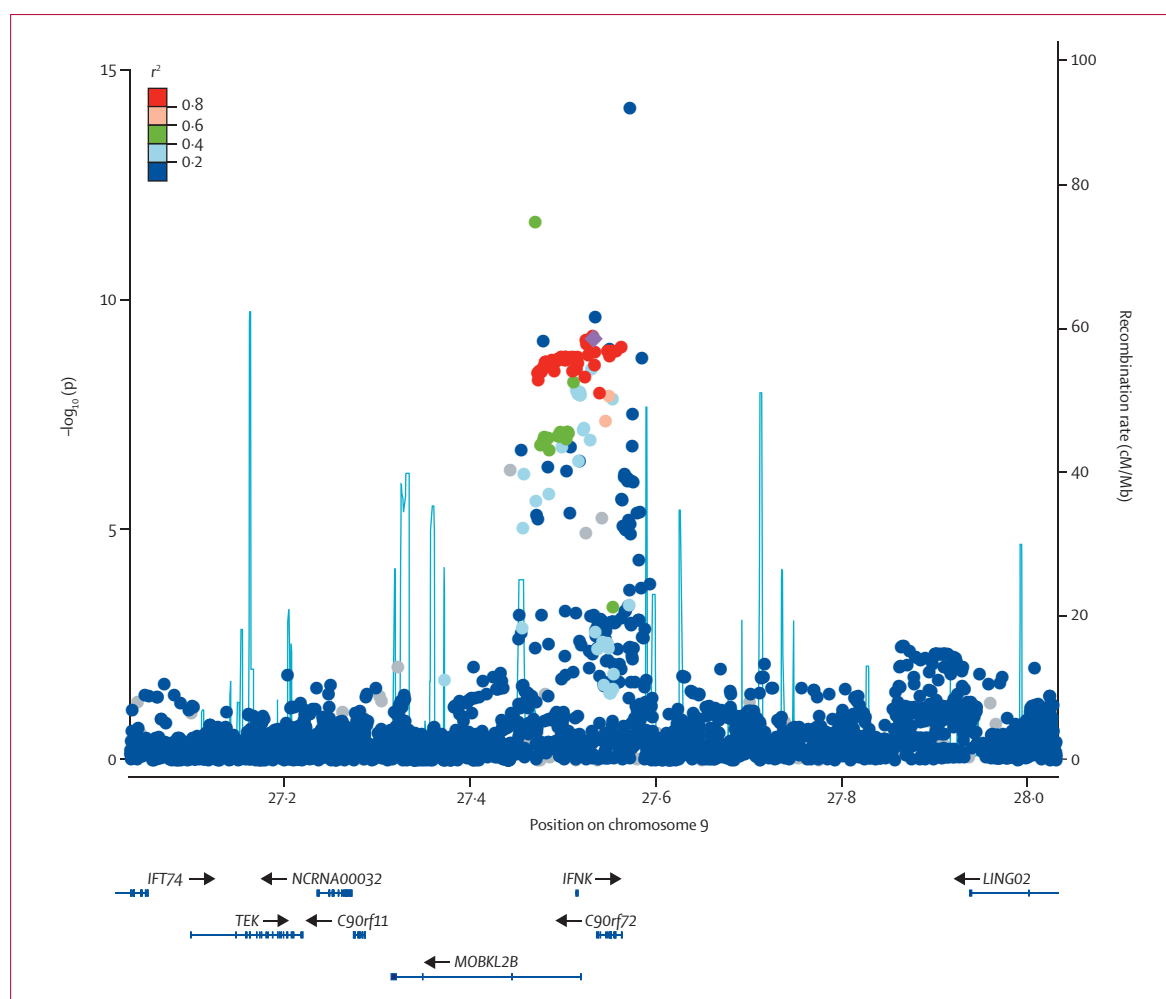


Figure 2: p values for association in the joint analysis



**Figure 3: Genetic architecture of the associated region**

Circles= $-\log_{10}$  of the p value of association with amyotrophic lateral sclerosis for typed and imputed single nucleotide polymorphisms in the joint analysis, coloured according to linkage disequilibrium with rs3849942 (shown as a purple diamond) as measured by  $r^2$ . Blue lines=recombination rate across the locus based on the 1000 genomes project. Genes in the region are displayed below the graph. Arrows indicate the direction of transcription.

*UNC13A* were the only significantly associated loci. The association of *UNC13A* with ALS was weaker in the joint analysis than in the study that identified the association,<sup>8</sup> which was well powered and replicated the results within the same study. Our independent samples do not show the association, possibly because the *UNC13A* variation is not a major cause of ALS in the UK, and so adding in these samples to the joint analysis we inevitably weaken the previously reported association. Given that the heritability of sporadic ALS is 0.61 (95% CI 0.38–0.78),<sup>26</sup> our findings suggest that even larger GWAS are required to have sufficient power to detect genetic variations responsible for ALS. A potential limitation of this study is that not all control individuals were age and sex matched, although they were matched for geographical origin and ancestry. For a disease as rare as ALS, age and sex matching does not increase statistical power, whereas it would for a more common disorder.

The associated 9p21 locus is the focus of much interest for researchers of both ALS and frontotemporal dementia. Although frontotemporal dementia and ALS are independent diseases, there is some overlap. About 5% of people with ALS have overt frontotemporal dementia and up to 51% have subtle cognitive deficits that suggest frontal and temporal lobe dysfunction.<sup>27</sup> Conversely, about 50% of people with frontotemporal dementia have features of motor neuron degeneration.<sup>28</sup> A locus on chromosome 9 has been independently linked to ALS–frontotemporal dementia in seven families;<sup>2–7</sup> this 3.6 Mb locus is defined across studies by the flanking markers D9S169 and D9S251. The SNPs we have identified lie within this region, with the peak association at 106.5 Kb. A GWAS that used pathological subtyping of patients with frontotemporal dementia to increase homogeneity also identified the same SNPs as those most strongly associated with sporadic

frontotemporal dementia (rs3849942  $p=1.38 \times 10^{-6}$  and rs2814707  $p=5.72 \times 10^{-6}$ ).<sup>29</sup> Therefore, variation in the same gene is likely to be responsible for sporadic ALS, sporadic frontotemporal dementia, and familial ALS–frontotemporal dementia. Assuming that the association locus is the same as the linkage locus, the region of interest in families with ALS–frontotemporal dementia must now be thought to be much narrower than that defined by recombination; however, if a synthetic association of rare variants and ALS is being detected, the disease-causing variation could still lie a substantial distance from the association peak.<sup>30</sup>

Penetrance is about 20% in the family with the strongest association between chromosome 9 and ALS–frontotemporal dementia.<sup>3</sup> Thus, familial variants might underlie at least some of the sporadic cases of ALS and frontotemporal dementia. The large signal seen in our independent study of UK samples is consistent with a founder effect in the UK population, which possibly extends out to other populations. In this respect, that a GWAS of a cohort of Finnish patients with ALS and controls also detected the chromosome 9p locus (rs3849942  $p=9.11 \times 10^{-11}$ ) is interesting.<sup>31</sup> This association signal was driven mainly by people with familial ALS; the most strongly associated SNP in this subset of patients was rs2225389 ( $p=2.23 \times 10^{-12}$ ). All the patients with familial ALS who were not homozygous for the D90A allele of the *SOD1* gene, and almost 20% of the patients with sporadic ALS, shared a 42-SNP haplotype,

which strongly suggests a founder effect for the chromosome 9p locus, at least in Finland. However, the penetrance of the associated risk allele of rs3849942 in sporadic ALS in our study is only slightly higher than the background risk of ALS (0.36–0.45%),<sup>1</sup> whereas the penetrance of the familial variants is about 20%; thus, the likelihood that the same variation is responsible for both familial and sporadic ALS is questionable, which makes a founder effect less likely. The significant association of rs3849942 with sporadic ALS in our joint analysis including multiple populations suggests that a common genetic variation might be contributing to disease risk. The two-SNP haplotype that was most strongly associated with ALS in the joint analysis (rs3849942-rs10122902) mimics the behaviour of the *APOE* association with Alzheimer's disease; the three *APOE* alleles— $\epsilon 2$ ,  $\epsilon 3$ , and  $\epsilon 4$ —are in fact pairs of alleles at two SNPs forming a haplotype. When a haplotype is formed from alleles at two SNPs there are four possible haplotypes, although not all will necessarily exist in the population. For both the rs3849942-rs10122902 haplotype in ALS and the *APOE* haplotype in Alzheimer's disease, only three of the four possible haplotypes are present, and the fourth does not exist in the population. In both cases, all identified haplotypes are common, and one is associated with risk of the disease (*APOE*  $\epsilon 4$  in Alzheimer's disease; AG in ALS), one is associated with protection from the disease (*APOE*  $\epsilon 2$  in Alzheimer's disease; GG in ALS), and one is neutral (*APOE*  $\epsilon 3$  in Alzheimer's disease; GA in ALS). Thus, common variation, as with *APOE* in Alzheimer's disease, rather than a founder effect, could be responsible for sporadic ALS. However, because the penetrance at a functional variant might be much higher than at an associated tag SNP, which would account for the discrepancy between the penetrance for familial ALS and sporadic ALS, a founder effect cannot be excluded. There must be multiple variants worldwide because families with chromosome 9p-linked ALS–frontotemporal dementia do not all share a common haplotype, and so multiple founders must exist. However, the chromosome 9p locus is important in ALS because the population attributable risk estimate shows that it is involved in 10% of all cases of sporadic ALS.

The disease-associated SNPs we have identified lie between two genes, *C9orf72* and *MOBK2B*, and are close to *IFNK*, which lies in an intron of *MOBK2B*. The protein encoded by *MOBK2B* is Mps one binder kinase activator-like 2B, which is involved in kinase regulation. On the opposite strand at the same location is *IFNK*, which encodes interferon  $\kappa$  precursor, which is important for immunity to viral infection. The SNPs comprising the most strongly associated haplotype (rs3849942 and rs10122902) and imputed SNP (chr9:27573575) overlap or lie within *C9orf72*, which encodes a protein that might be involved in cell development and possibly spermatogenesis.<sup>32</sup>

#### Panel: Research in context

##### Systematic review

To identify previously published genome-wide association studies (GWAS) in amyotrophic lateral sclerosis (ALS), we searched PubMed and the ALSOD database (<http://alsod.iop.kcl.ac.uk>), and contacted known ALS genetics research groups. Four GWAS reported findings that reached genome-wide significance;<sup>8–12</sup> we designed the independent study to replicate these. Association studies were selected for the joint analysis if research diagnostic criteria for ALS were used, if microarrays of at least 300 000 single nucleotide polymorphisms were used, and if individual rather than pooled genotypes were available. Because the main finding of the independent and joint analyses was of an association to a region on chromosome 9 also known to be associated with familial ALS–frontotemporal dementia and sporadic frontotemporal dementia, we searched PubMed and contacted research groups and identified a further seven families with ALS–frontotemporal dementia<sup>2–4,7</sup> and a GWAS of frontotemporal dementia.<sup>29</sup> These linkage data have been well summarised in a linkage paper.<sup>4</sup>

##### Interpretation

The ALS-associated locus on chromosome 9p was independently replicated, and no other loci were identified, suggesting that even larger genome-wide association studies are needed. We have shown the association signal on chromosome 9p lies in a well circumscribed block of linkage disequilibrium of about 106.5 Kb in size, which is flanked by markers rs4879515 and rs702231 and includes three annotated genes. This locus has also previously been implicated in sporadic frontotemporal dementia and in families with ALS–frontotemporal dementia. Because there is pathological and clinical overlap of ALS and frontotemporal dementia, the most logical interpretation is that dysfunction of a single gene product is responsible for some cases of ALS, frontotemporal dementia, and ALS–frontotemporal dementia.



Although linkage of ALS–frontotemporal dementia to this region of chromosome 9 was first identified in 2006,<sup>23</sup> sequencing has not yet yielded a disease-causing variant such as a non-synonymous SNP or splice-site variation. If this remains the case, one explanation could be that the variation is tagged by SNPs but not easily detected by exome capture and high throughput sequencing or Sanger sequencing. The disease-causing variant might be in an unknown exon or gene and has therefore not been sequenced, or it might be a trinucleotide repeat expansion in an intron, although such repeats are not seen in the surrounding region. Copy number variation is not necessarily easily detected, depending on the sequencing method used, but three microarray-based studies in ALS have not detected ALS-associated copy number variation in this region of chromosome 9.<sup>33–35</sup> Finally, a microRNA or non-annotated gene that has not been examined, or an inversion, which would be difficult to detect by sequencing, might be the disease-causing variant. Our results suggest genetic variation at this locus on chromosome 9 causes sporadic amyotrophic lateral sclerosis and familial ALS–frontotemporal dementia. Resequencing studies and then functional analyses are needed to identify the defective gene.

#### Contributors

AS, KM, JH, RWO, and AAC did the genotyping and analyses. SN, MEW, BS, CV, LJ, OA, and JFP did the analyses. KM, JHV, MAVe, LHvB, WR, OH, AEF, CML, AWB, PMA, IF, VS, AC, BJT, JM, VM, JEL, PM, JDG, HP, PNL, RHB, RWO, KEM, PJS, CES, and AAC contributed samples or genotypes or both. KM, PJS, CES, and AAC conceived the study. All authors contributed to the writing of the manuscript.

#### Contributing UK MND DNA bank investigators

Queen Elizabeth Hospital and University of Birmingham Hub  
J McConville, V Patterson (Belfast City Hospital); C Young (Walton Centre for Neurology and Neurosurgery); K Talbot (Radcliffe Hospitals NHS Trust, John Radcliffe Hospital); I Ormerod (North Bristol NHS Trust, Southmead Hospital); King's College Hospital and King's College London Hub R W Orrell (Royal Free Hospital and National Hospital for Neurology and Neurosurgery); C Hillier (Poole NHS Trust); J Frankel (Southampton University Hospitals NHS Trust, Southampton General Hospital); A Radunovic, A Malaspina (Barking, Havering and Redbridge University Hospitals NHS Trust); A Radunovic (Barts and the London NHS Trust); C Allen (Cambridge University Hospitals NHS Trust, Addenbrooks Hospital); J Zajicek (Derriford Hospital Plymouth).  
Royal Hallamshire Hospital and University of Sheffield Hub A J Wills (Nottingham University Hospitals NHS Trust, Queen's Medical Centre); R J Swinger (NHS Tayside, Ninewells Hospital); T L Williams (Newcastle Upon Tyne Hospitals NHS Foundation Trust, Royal Victoria Infirmary); D Mitchell (Lancashire Teaching Hospitals NHS Trust, Royal Preston Hospital); J Ealing (Greater Manchester Centre for Clinical Neuroscience and Hope Hospital Salford).

#### Conflicts of interest

WR has submitted a patent for a therapeutic molecule in neurodegeneration. OH has done consultancy work for Allergan, Merck, and Serono. VS has done consultancy work for Acceleron Pharma. AAC and PNL were funded by Trophos for a clinical trial of a therapeutic drug in ALS. PNL has done consultancy work for NeuroNova, Allon, and GlaxoSmithKline. RHB has done consultancy work for Biogen Idec and holds stock or stock options in Link Medicine and AviTx. PJS has submitted two patents for therapeutic molecules in ALS and has been a consultant to Sanofi-Aventis. All other authors declare no conflicts of interest.

#### Acknowledgments

The independent study case samples were obtained from the UK DNA bank for Motor Neuron Disease Research, which was funded by the Motor Neurone Disease Association grant 3/3 and the Wellcome Trust grant 070122/A/02/Z, with additional support from the National Institute for Health Research Dementias and Neurodegenerative Diseases Research Network (DeNDroN). Genotyping of the independent study case samples was funded by grants from the Angel Fund and ALS Therapy Alliance, with additional funding to RO and JH from the Motor Neurone Disease Association for genotyping (grant 10/57). We acknowledge use of genotype data from the British 1958 birth cohort DNA collection, which was funded by the Medical Research Council grant G0000934 and the Wellcome Trust grant 068545/Z/02. Controls for the DeCC study and BACCS were genotyped under the Medical Research Council grant G0701420. This project has been supported by the Prinses Beatrix Fonds, VSB Fonds, The Kersten Foundation, The Netherlands ALS Foundation, and The Adessium Foundation (LHvB). JHV is supported by grants from The Brain Foundation of the Netherlands. WR is supported by grants from the University of Leuven (Methusalem) and the Interuniversity Attraction Poles (IUAP) program P6/43 of the Belgian Federal Science Policy Office. VS is supported by the Italian Ministry of Health (Ricerca Finalizzata 2007 number 31). This work was supported in part by the Intramural Research Program of the National Institutes of Health, and the National Institute on Aging (Z01-AG000949-02). AC is supported by the Italian Ministry of Health (Ricerca Sanitaria Finalizzata, 2007, number 30), the Fondazione Vialli e Mauro per la SLA (grant 3/2008), and the Federazione Italiana Giuoco Calcio (Grant 2/2008). We thank the ALS Association, the ALS Society of Canada, MND Iceland, and the National Institute for Health Research Biomedical Research Centre for Mental Health for support, and Kuang Lin for programming advice.

#### References

- Johnston CA, Stanton BR, Turner MR, et al. Amyotrophic lateral sclerosis in an urban setting: a population based study of inner city London. *J Neurol* 2006; **253**: 1642–43.
- Morita M, Al-Chalabi A, Andersen PM, et al. A locus on chromosome 9p confers susceptibility to amyotrophic lateral sclerosis and frontotemporal dementia. *Neurology* 2006; **66**: 839–44.
- Vance C, Al-Chalabi A, Ruddy D, et al. Familial amyotrophic lateral sclerosis with frontotemporal dementia is linked to a locus on chromosome 9p13.2–21.3. *Brain* 2006; **129**: 868–76.
- Boxer AL, Mackenzie IR, Boeve BF, et al. Clinical, neuroimaging and neuropathological features of a new chromosome 9p-linked FTD-ALS family. *J Neurol Neurosurg Psychiatry* 2010; published online June 20. DOI:10.1136/jnnp.2009.204081.
- Valdmanis PN, Dupre N, Bouchard JP, et al. Three families with amyotrophic lateral sclerosis and frontotemporal dementia with evidence of linkage to chromosome 9p. *Arch Neurol* 2007; **64**: 240–45.
- Gijssels I, Engelborghs S, Maes G, et al. Identification of 2 loci at chromosomes 9 and 14 in a multiplex family with frontotemporal lobar degeneration and amyotrophic lateral sclerosis. *Arch Neurol* 2010; **67**: 606–16.
- Benajiba L, Le Ber I, Camuzat A, et al. TARDBP mutations in motoneuron disease with frontotemporal lobar degeneration. *Ann Neurol* 2009; **65**: 470–73.
- van Es MA, Veldink JH, Saris CG, et al. Genome-wide association study identifies 19p13.3 (UNC13A) and 9p21.2 as susceptibility loci for sporadic amyotrophic lateral sclerosis. *Nat Genet* 2009; **41**: 1083–87.
- Duncley T, Huentelman MJ, Craig DW, et al. Whole-genome analysis of sporadic amyotrophic lateral sclerosis. *N Engl J Med* 2007; **357**: 775–88.
- van Es MA, Van Vught PW, Blauw HM, et al. ITPR2 as a susceptibility gene in sporadic amyotrophic lateral sclerosis: a genome-wide association study. *Lancet Neurol* 2007; **6**: 869–77.
- Cronin S, Berger S, Ding J, et al. A genome-wide association study of sporadic ALS in a homogenous Irish population. *Hum Mol Genet* 2008; **17**: 768–74.
- van Es MA, van Vught PW, Blauw HM, et al. Genetic variation in DPP6 is associated with susceptibility to amyotrophic lateral sclerosis. *Nat Genet* 2008; **40**: 29–31.

- 13 Fogh I, D'Alfonso S, Gellera C, et al. No association of DPP6 with amyotrophic lateral sclerosis in an Italian population. *Neurobiol Aging* 2009; published online June 12. DOI:10.1016/j.neurobiolaging.2009.05.014.
- 14 Lewis CM, Ng MY, Butler AW, et al. Genome-wide association study of major recurrent depression in the UK population. *Am J Psychiatry* 2010; **167**: 949–57.
- 15 Gaysina D, Cohen-Woods S, Chow PC, et al. Association of the dystrobrevin binding protein 1 gene (DTNBP1) in a bipolar case-control study (BACCS). *Am J Med Genet B Neuropsychiatr Genet* 2009; **150B**: 836–44.
- 16 Landers JE, Melki J, Meininger V, et al. Reduced expression of the kinesin-associated protein 3 (KIFAP3) gene increases survival in sporadic amyotrophic lateral sclerosis. *Proc Natl Acad Sci USA* 2009; **106**: 9004–09.
- 17 Schymick JC, Scholz SW, Fung HC, et al. Genome-wide genotyping in amyotrophic lateral sclerosis and neurologically normal controls: first stage analysis and public release of data. *Lancet Neurol* 2007; **6**: 322–28.
- 18 Wroe R, Wai-Ling Butler A, Andersen PM, Powell JF, Al-Chalabi A. ALSOD: the amyotrophic lateral sclerosis online database. *Amyotroph Lateral Scler* 2008; **9**: 249–50.
- 19 Purcell S, Cherny SS, Sham PC. Genetic power calculator: design of linkage and association genetic mapping studies of complex traits. *Bioinformatics* 2003; **19**: 149–50.
- 20 Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006; **38**: 904–09.
- 21 Patterson N, Price AL, Reich D. Population structure and eigenanalysis. *PLoS Genet* 2006; **2**: e190.
- 22 Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; **81**: 559–75.
- 23 Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5**: e1000529.
- 24 International HapMap Consortium. The international HapMap project. *Nature* 2003; **426**: 789–96.
- 25 Marchini J, Howie B, Myers S, McVean G, Donnelly P. A new multipoint method for genome-wide association studies by imputation of genotypes. *Nat Genet* 2007; **39**: 906–13.
- 26 Al-Chalabi A, Fang F, Hanby MF, et al. An estimate of amyotrophic lateral sclerosis heritability using twin data. *J Neurol Neurosurg Psychiatry* (in press).
- 27 Ringholz GM, Appel SH, Bradshaw M, Cooke NA, Mosnik DM, Schulz PE. Prevalence and patterns of cognitive impairment in sporadic ALS. *Neurology* 2005; **65**: 586–90.
- 28 Lomen-Hoerth C. Characterization of amyotrophic lateral sclerosis and frontotemporal dementia. *Dement Geriatr Cogn Disord* 2004; **17**: 337–41.
- 29 Van Deerlin VM, Sleiman PM, Martinez-Lage M, et al. Common variants at 7p21 are associated with frontotemporal lobar degeneration with TDP-43 inclusions. *Nat Genet* 2010; **42**: 234–39.
- 30 Dickson SP, Wang K, Krantz I, Hakonarson H, Goldstein DB. Rare variants create synthetic genome-wide associations. *PLoS Biol* 2010; **8**: e1000294.
- 31 Laaksovirta H, Peuralinna T, Schymick JC, et al. Chromosome 9p21 in amyotrophic lateral sclerosis in Finland: a genome-wide association study. *Lancet Neurol* 2010, published online Aug 31. DOI:10.1016/S1474-4422(10)70184-8.
- 32 Beaver JE, Tasan M, Gibbons FD, Tian W, Hughes TR, Roth FP. FuncBase: a resource for quantitative gene function annotation. *Bioinformatics* 2010; **26**: 1806–07.
- 33 Wain LV, Pedrosa I, Landers JE, et al. The role of copy number variation in susceptibility to amyotrophic lateral sclerosis: genome-wide association study and comparison with published loci. *PLoS ONE* 2009; **4**: e8175.
- 34 Cronin S, Blauw HM, Veldink JH, et al. Analysis of genome-wide copy number variation in Irish and Dutch amyotrophic lateral sclerosis populations. *Hum Mol Genet* 2008; **17**: 3392–98.
- 35 Blauw HM, Veldink JH, van Es MA, et al. Copy-number variation in sporadic amyotrophic lateral sclerosis: a genome-wide screen. *Lancet Neurol* 2008; **7**: 319–26.